

PAPER • OPEN ACCESS

## A two stage data envelopment analysis model with undesirable output

To cite this article: Adam Shariff Adli Aminuddin *et al* 2017 *J. Phys.: Conf. Ser.* **890** 012110

View the [article online](#) for updates and enhancements.

### Related content

- [Data envelopment analysis with upper bound on output to measure efficiency performance of departments in Malaikulsaleh University](#)  
Dahlan Abdullah, Saib Suwilo, Tulus et al.
- [Selection Input Output by Restriction Using DEA Models Based on a Fuzzy Delphi Approach and Expert Information](#)  
Roslah Arsad, Mohammad Nasir Abdullah, Suriana Alias et al.
- [The application of DEA model in enterprise environmental performance auditing](#)  
F Li, L Y Zhu, J D Zhang et al.

# A two stage data envelopment analysis model with undesirable output

Adam Shariff Adli Aminuddin<sup>1</sup>, Nur Izzati Jaini<sup>1</sup>,  
Maznah Mat Kasim<sup>2</sup> and Mohd Kamal Mohd Nawawi<sup>2</sup>

<sup>1</sup>Faculty of Industrial Sciences and Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Kuantan, Pahang, Malaysia

<sup>2</sup>School of Quantitative Sciences, College of Arts & Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

E-mail: adamshariff@ump.edu.my

**Abstract.** The dependent relationship among the decision making units (DMU) is usually assumed to be non-existent in the development of Data Envelopment Analysis (DEA) model. The dependency can be represented by the multi-stage DEA model, where the outputs from the precedent stage will be the inputs for the latter stage. The multi-stage DEA model evaluate both the efficiency score for each stages and the overall efficiency of the whole process. The existing multi stage DEA models do not focus on the integration with the undesirable output, in which the higher input will generate lower output unlike the normal desirable output. This research attempts to address the inclusion of such undesirable output and investigate the theoretical implication and potential application towards the development of multi-stage DEA model.

## 1. Introduction

Data envelopment analysis (DEA) is a method introduced by Charnes, Cooper and Rhodes in 1978 to determine the relative efficiency of decision making units (DMU), comprises of input and output. Since then, various seminal papers have been introduced to further expand the model and numerous application has proved the effectiveness of DEA [1]. DEA is decent in handling multi indicator model but it is also subject to some limitations. A general DEA model assumes the DMUs to be independent, homogenous and [2] even addressed that the complex interaction issues between DMU remains to be an open problem.

Several attempts have been suggested to consider the internal structure between DMUs which includes shared flow, multi-level, and network models which include dynamic and multi-stage representations [3]. The most recent research in the area of NS-DEA model investigated this problem from the perspective of multi-stage model in which the DMUs is commonly separated into two stages. Through this approach, the DMU is being measured separately according to the importance of the indicator before aggregating it as a central efficiency. [4] and [5] suggested that the output of an indicator in multi-stage model to become the input for another indicator to represent the dependency relation among indicators. This form of dependency actually combines the indicators and treat it as a single entity rather than separate inter dependent indicators, which is impractical for some instances. For example, the Gross Domestic Product (GDP) has some influence towards employment rate, but if GDP is low, it does not mean the employment rate will definitely be low.

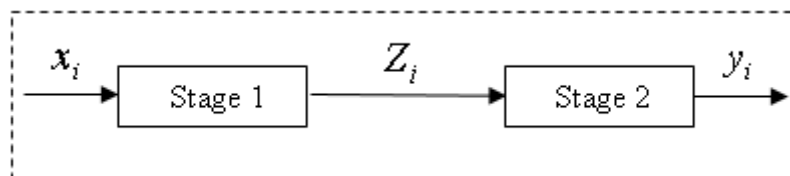


A conventional DEA model also behaves under the assumption that the input which provides higher output to be more efficient which includes for example profit, income, and total number of productions. In spite of that, there are many real world cases which needs to be modelled in which the lower output indicates higher efficiency. The examples of this outputs includes environmental pollution emission, waiting time, inventory size and total usage time. Therefore, an input may produce two types of outputs which are good output and bad output, which can be modelled as separable or non-separable (NS-DEA) [6]. The difference between them is that in the NS-DEA model, the changes of good output will affect the production of bad output and vice versa. The post aggregation and multi stage methods does not necessarily provide accurate representation of efficiency score especially when the outcome was not validated and undesirable output is not considered. Based on this limitation, we attempt to address the multi-stage data envelopment analysis model in two stages with the inclusion of undesirable output.

## 2. Interaction between two stages

From the perspective of multi stage model, the internal structure of DMU consists of sequential consecutive stages [7] as shown in Figure 1. A basic interaction or intermediate measures,  $Z_i$  between two stages maybe assumed to be dependent, in which the input of Stage 1 will produce an output, which will become the input for Stage 2 to produce an aggregated output. In spite of that,  $Z_i$  may also be independent and there are several types of  $Z_i$  suggested by previous researches to calculate the overall efficiency score,  $\theta$  as exhibited in Table 1.

All of the previous researches mentioned in Table 1 did not consider the inclusion of undesirable output. Thus, this research attempts to advocate the undesirable output point of view. As a starting point, this research examines the consecutive two stage model based on the undesirable output model by [8].



**Figure 1.** Basic two stage interaction in DMU.

**Table 1.** Types of intermediate measure,  $Z_i$ .

Intermediate Measure, $Z_i$	Description
Consecutive [7]	$\theta_1 = x_2$ $\theta$ is determined using conventional DEA model
Multiplicative [9]	$\theta = \theta_1 \cdot \theta_2$
Additive [10]	$\theta = t_1 \cdot \theta_1 + t_2 \cdot \theta_2$ s.t. $t_1 + t_2 = 1$
Alternative Additive 1 [11]	$\theta = \frac{1}{2} \cdot \theta_1 + \frac{1}{2} \cdot \theta_2$
Alternative Additive 2 [11]	$w_1 \cdot \theta_1 + w_2 \cdot \theta_2$
Notation:	
$\theta$ = Total efficiency score	
$\theta_1$ = Efficiency score Stage 1	
$\theta_2$ = Efficiency score Stage 2	
$x_2$ = Input Stage 2	
$t_1$ = Weight 1	
$t_2$ = Weight 2	
$w_1$ = User defined weight 1	
$w_2$ = User defined weight 2	

### 3. Methodology

Given the undesirable output model by [8],

$$\begin{aligned}
 & \text{Max } \beta \\
 \text{s.t.} \quad & X\lambda \leq x_o \\
 & Y^d\lambda \geq y_o^d + \beta y_o^d \\
 & Y^u\lambda \leq y_o^u - \beta y_o^u \\
 & \max\{y_o^u\} \geq y_o^u - \beta y_o^u \\
 & \lambda \geq 0
 \end{aligned} \tag{1}$$

Let input  $x$ , and output  $y$  consists of desirable and undesirable outputs  $y = (y^d, y^u)$ , where  $\theta^* = \{(x, y^d, y^u) | x \geq X\lambda, y^d \leq Y\lambda, y^u \geq 0\}$  under constant return to scale (CRS). The optimal solution of (1) is given by  $\beta^*$ . If  $\beta^* = 0$ , where  $\lambda_0 = 1, \lambda_j = 0 (j \neq 0)$ , then the DMU is efficient. Otherwise  $\beta^* > 0$  implies that the DMU is not efficient. In this research, the values of  $\beta_1$  was made to be the input for Stage 2, and the overall efficiency,  $\theta$  was determined using the conventional envelopment form CCR model.

The calculation of undesirable output model of (1) for Stage 1 is done using MATLAB DEA solver by [8]. The total number of DMUs is fixed at 5, the input for and the portion of desirable and undesirable output for Stage 1 is based on data by [8] for result comparison purpose. The normalized input and output data for Stage 2 is based on simulated data. The data used is also assumed to be separable, and the desirable output do not have any influence with undesirable output, and vice versa. Then, the values of normalized  $\beta$ ,  $\beta^* = 1 - \beta$  becomes the actual input for Stage 2,  $X_2$  which corresponds to the normalized output for Stage 2,  $Y_2^*$ . The overall efficiency,  $\theta$  was determined using the conventional CCR model to comply with CRS assumption of (1), with the assistance of Win4DEAP2 software. The summary of input data is shown in Table 2.

**Table 2.** Two stage model input.

DMU	Stage 1			Stage 2	
	$X_1$	$Y_1^d$	$Y_1^u$	$X_2$	$Y_2$
A	1.000	7.000	2.000	1.000	5.000
B	1.000	5.000	5.000	1.000	3.000
C	1.000	1.000	3.000	1.000	9.000
D	1.000	3.000	3.000	1.000	1.000
E	1.000	4.000	2.000	1.000	7.000

**4. Results and discussions**

The final efficiency score,  $\theta$  of two stage undesirable output data is presented in Table 3. In addition, the efficiency score and ranking of this research approach which is based on basic consecutive  $Z_i$ , is compared with the separate Stage 1 and Stage 2 CCR models, given by  $\theta_1$  and  $\theta_2$ , multiplicative  $Z_i$  and additive  $Z_i$ . The additive weight for additive  $Z_i$ ,  $t_1$  and  $t_2$  is being set equally to 0.5 for each stage similar to [10] and [11] as displayed in Table 4 and Table 5.

**Table 3.** Two stage model output.

DMU	Stage 1		Stage 2	
	$\beta$	$\beta^* = X_2$	$Y_2^*$	$\theta_{CCR}$
A	0.000	1.000	5.000	0.556
B	0.400	0.600	1.800	0.333
C	0.826	0.174	1.565	1.000
D	0.556	0.444	0.444	0.111
E	0.273	0.727	5.092	0.778

**Table 4.** DMU efficiency score between  $Z_i$ .

DMU	$\theta_1$	$\theta_2$	Consecutive*	Multiplicative	Additive ( $t = 0.5$ )
A	1.000	0.556	0.097	0.556	0.778
B	0.600	0.333	0.097	0.200	0.467
C	0.174	1.000	1.000	0.174	0.588
D	0.444	0.111	0.044	0.049	0.278
E	0.727	0.778	0.186	0.566	0.753

Based on Table 3, the undesirable output of Stage 1 indicates that DMU A to be efficient and four other DMUs to be not efficient. However, after two consecutive stages, results indicates that DMU C is the efficient DMU, whereas four other DMUs is not efficient. The comparative analysis of ranking in Table 5 shows that the output of consecutive  $Z_i$  is almost similar with  $\theta_2$ . Therefore, it is possible that the efficiency of consecutive  $Z_i$  is highly affected by the DMUS's performance in Stage 2 despite the performance in Stage 1. The multiplicative  $Z_i$  is not consistent with  $\theta_1$ ,  $\theta_2$  as well as consecutive and additive  $Z_i$ . Lastly, the additive  $Z_i$  ranking is not similar with other  $Z_i$ , with fair margin difference between  $\theta_1$  and  $\theta_2$  for DMU B, C, D, and E, with DMU A is bias towards  $\theta_1$ .

**Table 5.** DMU efficiency ranking between  $Z_i$ .

DMU	$\theta_1$	$\theta_2$	Consecutive*	Multiplicative	Additive ( $t = 0.5$ )
A	1	3	3	2	1
B	3	4	3	3	4
C	5	1	1	4	3
D	4	5	5	5	5
E	2	2	2	1	2

## 5. Conclusion and future directions

This research demonstrates the utilization of multi stage DEA model with undesirable output which stems from consecutive  $Z_i$ . The result is also compared with multiplicative and additive  $Z_i$ . This research concludes that additive  $Z_i$  performs the best based on this research's data set and conversely, multiplicative  $Z_i$  performs negatively. The consecutive  $Z_i$  also highly affected by the latter stage which may suggest that this type of interaction will not accurate in representing the overall  $\theta$ , especially if it involves  $n$ -stages. For future research, it is suggested that the additive model is tested under various values of  $t$ . Besides, the model should be expanded for variable return to scale (VRS) and  $n$ -stages model. Finally, it is worth to further investigate undesirable output model as non-separable and applied into other internal structure forms of shared flow, multi-level and dynamic network model.

## Acknowledgments

The authors would like to thank the Malaysia Ministry of Higher Education (MoHE) for the financial support under the Fundamental Research Grant Scheme (FRGS).

## References

- [1] Cooper W W, Seiford L M and Zhu J 2011 *Data envelopment analysis: History, models, and interpretations* In *Handbook on data envelopment analysis* (Springer US)
- [2] Despotis D K, Sotiros D and Koronakos G 2016 *Omega* **61** pp 35-48
- [3] Castelli L, Pesenti R and Ukovich W 2010 *Annals of Operations Research* **173**(1) pp 207-235
- [4] Izadikhah M and Saen R F 2016 *Transportation Research Part D: Transport and Environment* **49** pp 110-126
- [5] Galan-Martin A, Guillen-Gosalbez G, Stamford L and Azapagic A 2016 *Computers & Chemical Engineering* **90** pp 188-200
- [6] Cooper W W, Seiford L M and Tone K 2007 *Data envelopment analysis: A comprehensive text with modes, applications, references and DEA-solver software* (Springer Science & Business Media)
- [7] Seiford L M and Zhu J 1999 *Management science* **45**(9) pp 1270-1288
- [8] Álvarez I, Barbero J and Zofio Prieto J L 2016 *A Data Envelopment Analysis Toolbox for MATLAB* (No. 2016/03) Universidad Autónoma de Madrid (Spain) Department of Economic Analysis (Economic Theory and Economic History)
- [9] Kao C and Hwang S N 2008 *European Journal of Operational Research* **185**(1) pp 418-429
- [10] Chen Y, Cook W D, Li N and Zhu J 2009 *European Journal of Operational Research* **196**(3) 1170-1176
- [11] Despotis D K and Koronakos G 2014 *Procedia Computer Science* **31** pp 299-307